

CUPRINS

INTRODUCERE	7
1. TEXT MINING – CARACTERISTICI ȘI DOMENII DE APLICABILITATE	9
1.1. Caracteristicile documentelor unei organizații	11
1.2. Text mining – domeniu interdisciplinar	16
1.3. Abordări în text mining	17
1.4. Domenii de aplicabilitate ale text mining	20
1.5. Arhitectura funcțională a unui sistem de text mining	21
2. REPREZENTAREA DOCUMENTELOR	26
2.1. Reprezentări descriptive	27
2.2. Caracteristicile documentelor	31
2.3. Reprezentarea documentelor în spațiul vectorial	34
2.4. Reprezentarea specifică documentelor web	37
3. CLASIFICAREA AUTOMATĂ A DOCUMENTELOR	45
3.1. Învățarea automată în clasificarea textelor	47
3.2. Aplicații ale clasificării textelor	49
3.3. Algoritmi de clasificare	51
3.3.1. <i>Clasificatori simbolici</i>	52
3.3.2. <i>Clasificatori liniari</i>	54
3.3.3. <i>Clasificatori bazați pe exemple</i>	64
3.3.4. <i>Combinarea algoritmilor de clasificare</i>	65
3.4. Evaluarea algoritmilor de clasificare a documentelor	67
3.4.1. <i>Precizia și relevanța</i>	67
3.4.2. <i>Eficiența</i>	70
3.4.3. <i>Utilitatea</i>	72
3.4.4. <i>Acuratețea și eroarea</i>	73
3.5. Compararea algoritmilor de clasificare a documentelor	73

4. WEBUL SEMANTIC – CARACTERISTICI ȘI LIMBAJE	75
4.1. Reprezentarea cunoștințelor	76
4.2. Resource description framework.....	81
4.3. Limbaje de descriere a ontologiilor.....	85
4.4. Ontologii de proces	88
4.5. Dezvoltarea ontologiilor prin text mining.....	93
5. REPREZENTAREA CUNOȘTINȚELOR IN DOCUMENTELE WEB	
ÎN SISTEMUL OntoDev	96
5.1. Arhitectura funcțională.....	97
5.2. Algoritmul de reprezentare a documentelor web	98
5.3. Schema logică a bazei de date.....	102
5.4. Reprezentarea conceptelor ontologiei academice în documentele web ...	104
5.5. Analiza logurilor	107
5.6. Identificarea modelelor de comportament al utilizatorilor de internet.....	113
6. CONCLUZII	124
BIBLIOGRAFIE.....	126